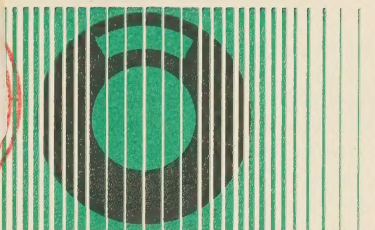
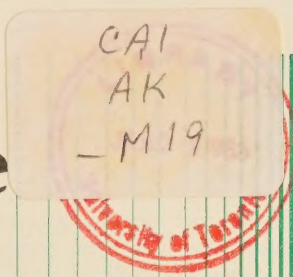




# Machine Readable Archives



## BULLETIN

ISSN 0821-3658

### New Acquisitions

The *Election Results, 1984* data file has been acquired from Elections Canada (formerly the Chief Electoral Office). This file provides poll by poll results from the September 1984 federal election. Data files for the federal elections in 1972, 1974, 1979 and 1980 are also available.

### EDP Records Scheduling Program

For several years the Public Archives of Canada (PAC) has depended, in large part, upon the scheduling function for the identification, appraisal, acquisition and preservation of historically valuable recorded information generated by federal government departments and agencies. Schedules are documents prepared by the records managers of federal agencies that describe the length of time groupings of related records are to be kept by an agency for operational purposes and the length of time they are to be kept by the Public Archives as dormant records in regional records centres. In order for archivists to appraise records for their archival value, agencies are required to submit these schedules to the Dominion Archivist for approval, at which time archivists are provided with an opportunity to indicate which records must be transferred to the Archives Branch of the PAC at the end of the retention periods indicated in the schedules.

The scheduling function has depended upon the establishment of a well designed records management function in each government agency. Recently, the strength of this function was reinforced by the passage of Access to Information and Privacy legislation and the issuance of a government-wide Treasury Board policy on records management. Although the scheduling function has worked reasonably well for hard-copy or textual records, it has not enhanced to any great extent information recorded in other forms or information maintained by organizational areas that were traditionally separate from the records management function. Machine readable data managed in computerized systems were rarely accounted for by the records management areas of most agencies. Although attempts to identify machine readable records held by government agencies were made in the seventies, it was not until the creation of the EDP Information Systems Section in the Machine Readable Archives (MRA) that a systematic program to identify, describe and

schedule data managed in computerized systems was developed. Recognizing the fact that individuals from several areas within an agency must be involved in the identification and scheduling of EDP records, a project team approach has been recommended. In addition to PAC representatives, a project team should be composed of the information systems representative, the records managers, the Access to Information and Privacy Coordinator, and those others who may be directly involved with the use of the data. The project team approach was tested in three pilot projects undertaken over the past year. The procedures for describing the data managed by automated systems (system overview) were tested and the results of the pilot projects were encouraging. Over the next few years it is hoped that those procedures will be used in all federal agencies.

#### Schedules for EDP Records: The System Overview Approach

The development of a schedule is based on the preparation of "system overviews." An overview is designed to provide a generic picture of the information content of the data managed by the system and a large scale map of the major stages through which data flow as they are processed by the system. As a result, agencies are not required to prepare exhaustive descriptions of all of their tape and disk holdings. The intention is to identify the computerized applications supported by the agency and, based on a top-down approach, to describe the key components that can benefit from the services offered by the Public Archives.

Many systems are designed to support only one function or application (e.g., a one-time survey, issuance of labels, etc.). Others may support multiple applications, each designed to perform a unique function often based on a common pool of data (e.g., a personnel application and a financial application supported by a common database of integrated personnel and financial data). As a result, a system overview may comprise one or several application descriptions.

The elements of the system overview, other than the administrative details such as name of the institution and organization unit, are as follows:

**System Title:** A descriptive, meaningful title of the automated system.

**Application Title:** A description of the name of each application supported by the system (a system may support one or, in the case of databases, multiple applications).

**Purpose:** A description of the purpose of the application, particularly in terms of the function or related functions it is designed to perform.

**Information Content:** A summary of the information content of the data processed by the application.

**Inputs/Processes/Outputs:** An overview of the major stages through which data are processed by the application. This description is not restricted to the initial inputs and final outputs. It is based on a system or data flow chart to provide a narrative or "story" of how data are processed by the system. The frequency with which certain steps are performed is also indicated.

**Documentation/Cross References:** A description of all the key documents (and their location) that describe the specifications, procedures, codes and other supporting information related to the application.

**Software Environment:** A generic reference to the software that is used to manage the data within the application.

**Hardware Environment:** A description of the hardware used to support the application.

**System Milestones:** A summary of the major events that have affected the design of the system (production year, years in which major system conversions and changes in record formats occurred, etc.).

**User and Technical Contacts:** Names, positions and titles of the major users and individuals responsible for the technical aspects of the application.

**Subject File Cross Reference:** A reference to the appropriate title and number within the agency subject file system to which the information processed by the application is indicated. The subject file system is maintained by the agency's records manager.

**Class of Records Cross Reference:** The name and registration number of the class of records to which the information managed by the application is related. This is a requirement of the new Access to Information legislation.

**Personal Information Bank/Class Cross Reference:** The name and registration number of the Personal Information Bank (if applicable) to which the personal information managed by





## "EDP Records..."

cont'd from page 1

the application is related. This is a requirement of the new Privacy Act.

*Statement of Retention and Disposal:* The length of time that the information, regardless of the physical storage media, is required to be maintained in accordance with the legislative or policy requirements (e.g., Privacy Act, Financial Administration Act, government-wide policies and regulations, etc.) and in accordance with its operational usefulness to the agency.

*Retention and Disposal of Physical Recording Media:* Identification of the key items (e.g., process files, input forms, etc.) associated with the administration of the system; and retention periods for each of these items based on the system specifications and consultation with records managers, users and PAC staff regarding the archival value of various items, the items that should be stored in federal records centres, and the confirmation that retention periods comply with legal or policy requirements.

## Link to Archival Appraisal

The system overviews are of great benefit to the archivist in undertaking the appraisal of the records. The "purpose" statement places the system in the context of a program activity, thus

providing the archivist with the organizational information required to establish the evidential value of the data. The "information content" description provides the archivist with an opportunity to appraise the informational value of the data.

The "inputs/processes/outputs" information, by describing the steps through which the data flow as they are processed by the system application, provides the archivist with the map required to determine which form or forms of recorded information should be acquired. The "production dates" provide the archivist with an indication of the number of years of data that could be involved in a potential acquisition. The "documentation/cross references" provide the archivist with a convenient finding aid to the documentation that is required not only for a detailed understanding of the data, but also to ensure the continued accessibility of those data that are designated as having archival value. The "software and hardware environment" description provides information concerning the technical requirements of the data, and the "system milestones" will provide details on the number of years of data plus important format changes.

The "user and technical contacts" provide the archivist with the names of those individuals who have the expertise to advise on both the intellectual and technical aspects of the system and who would generally assist in the appraisal and acquisition process.

The "subject file cross reference" provides the archivist with an opportunity to perform a comprehensive appraisal of all recorded information associated with the system. The "retention and disposal" information as described in both the overview and in the related area of the subject file system provides the archivist with an indication of the length of time that various forms of recorded information associated with the system (e.g., correspondence, questionnaires, transaction files, end-of-year history files) will be retained by the agency, and an opportunity, through the appraisal process, to preserve, beyond the end of their retention periods, those data appraised as having archival value.

This approach is designed to ensure that all forms of recorded information are accounted for in one comprehensive descriptive and scheduling process. By accounting for all recorded information associated with a given system, archivists need not be concerned that they are viewing only one slice of the system. As a result, all recorded information — both hardcopy and computerized — can be viewed in their proper context within a given system.

*This article is an abridged version of a paper by John McDonald, "An Approach to the Identification and Scheduling of EDP Data," presented at the Society of American Archivists, September 2, 1984, in Washington, D.C.*

## News and Notes

### 1985 IFDO/IASSIST Conference

The 1985 IFDO/IASSIST conference was held in Amsterdam, May 20-24. Approximately one hundred forty data archivists, researchers, administrators and librarians from twenty different countries participated in the sessions. The theme of the conference, "Public Access to Public Data," provided a forum in which to discuss such problems as identifying existing data; improving access to these data; identifying barriers that prevent access; improving the dissemination of information about data files; storage media; as well as privacy issues. A list of titles of conference sessions is provided. Pre-conference abstracts of papers are available through the Steinmetz Archives in Amsterdam. Efforts are being made to publish the papers presented at the conference.

#### Session Topics:

Public Data for Policy and Research  
Overcoming Barriers to Public Access  
Archive Law: Preservation, Access, Responsibility  
Statistical Techniques for Preserving and Evaluating the Confidentiality of Data Releases  
Issues of Privacy and Access  
Historical Public Data  
New Media and Technology  
Comprehensive On-line Statistical Information Systems  
New Dimensions in Data Storage and Use  
Cooperation between Producers, Distributors and Users  
Varied Uses of Public Data for Policy, Research and Instruction  
Techniques for Secondary Analysis  
Measurement of Social Class in Public Data

Methodological Issues in Using Public Data  
The Changing Census and the Search for New Sources: Problems and Solutions  
The Public's Need for Data  
Longitudinal Panel Studies  
Alternative Sources of Public Data

### IASSIST '86

Next year's conference will be held in Marina Del Rey (near Los Angeles), California, May 21-25. For information about the conference contact Jackie McGee, RAND Corporation, 1700 Main St., Santa Monica, California.

The articles in this issue were prepared by Sue Gavrel, MRA.



*Délais de conservation et d'élimination des données* : Période de conservation des données quel que soit le support utilisé, compte tenu de lois ou de politiques particulières (p. ex. Loi sur la protection des renseignements personnels, Loi sur l'administration financière, politiques et règlements du gouvernement) et de l'utilité réelle de l'information pour l'organisme.

*Conservation et élimination des supports d'enregistrement* : Identification des éléments-clés (p. ex. fichiers de traitement, formules d'intro-duction des données, etc.) associées à l'adminis-tration du système, établissement des délais de conservation de chacun des éléments d'après les spécifications et les discussions avec les archi-vistes de documents, les utilisateurs et les archi-vistes relativement à la valeur archivistique de certains documents, à la nécessité de faire appel aux centres de conservation aux lois et politiques en vigueur.

## Utilité de la vue d'ensemble pour l'archiviste

L'archiviste qui doit évaluer la teneur archi-vistique ou documentaire d'un document peut

## Notes et nouvelles

### IASSIST/IFDO 1985

Environ 140 archivistes, chercheurs, admi-nistrateurs et bibliothécaires d'une vingtaine de pays ont assisté à la conférence de l'IASSIST et de l'IFDO 1985 qui a eu lieu à Amsterdam du 20 au 24 mai. Le thème de la conférence « Accès public aux données publiques » s'est articulé autour des sujets suivants : l'identification des données existantes, la façon d'améliorer l'accès à ces données, les obstacles qui gênent l'accès à l'information, la façon d'améliorer la diffusion d'informations sur les fichiers de données, les supports de conservation, et la protection des renseignements personnels. Les ateliers sont énu-mérés ci-dessous. On peut obtenir des résumés d'exposés antérieurs à la conférence auprès des Archives Steimmetz à Amsterdam. Des démar-ches sont en cours en vue de publier les com-muniqués présentés à la conférence.

### Ateliers

Données publiques relatives aux politiques et à la recherche

Comment surmonter les obstacles qui gênent l'accès public  
Loi sur les archives : conservation, accès et responsabilités  
Statistique appliquée à la conservation du caractère confidentiel des données  
divulguées  
Accès à l'information et protection des renseignements personnels  
Données historiques  
Nouveaux médias et nouvelles techniques  
Systèmes d'information statistique en direct  
Nouvelles dimensions en matière de mémorisation et d'utilisation des données  
Collaboration entre producteurs, diffuseurs et utilisateurs  
Applications des données publiques : élaboration de politiques, recherche et enseignement  
Techniques d'analyse secondaire  
Répartition des classes sociales dans les données publiques

La DAO.

Les articles ont été rédigés par Sue Gavrel de la DAO.

### IASSIST 86

Méthodologies propres à l'utilisation des données publiques  
Modifications du recensement et recherche de nouvelles sources : problèmes et solutions  
Le public et son besoin d'accès aux données  
Enquêtes longitudinales  
Autres sources de données publiques  
La conférence de l'an prochain aura lieu à Marina Del Rey (près de Los Angeles) en Californie, du 21 au 25 mai. Pour tous ren-seignements, communiquer avec Jackie McGee, RAND Corporation, 1700 Main St., Santa Monica, Californie.

tirer parti de la vue d'ensemble du système. En effet, la rubrique « but » situe le système dans un contexte spécifique, et l'archiviste est des lors capable d'établir la valeur de témoignage des données. La rubrique « contenu de l'informa-tion » lui permet ensuite de déterminer la valeur informative de ces données.  
La rubrique « entrées/traitement/sorties » fournit à l'archiviste un organigramme qui le renseigne sur les supports d'information à acquérir. Les dates de production lui révèlent le nombre d'années sur lesquelles s'échelonnent les données. La rubrique « documentation/ren-vois » permet à l'archiviste de bien comprendre les données et de garantir l'accès à celles qui ont une valeur archivistique. La description du logi-ciel et du matériel fournit des renseignements de nature technique, tandis que la rubrique « événements marquants » renseigne l'archiviste sur le nombre d'années en cause et sur les grandes modifications aux structures d'enre-gistrement.  
La rubrique « utilisateurs et conseillers techniques » révèle à l'archiviste l'identité des personnes qui peuvent l'éclairer sur les proprié-tés du système et le seconder dans les processus d'évaluation et d'acquisition.  
Les renvois aux systèmes de classification permettent à l'archiviste d'effectuer une évalua-tion globale de toute l'information liée au sys-tème. Les données sur les délais de conservation

et d'élimination comprises dans la vue d'en-semble du système informatique et dans la par-tie correspondante du système de classification (p. ex. correspondance, questionnaires, dossiers de transactions, récapitulations annuelles de dos-siers) et, par le biais du processus d'évaluation, lui permettent de sauvegarder les données importantes une fois les délais de conservation écoulés.  
Grâce à cette méthode, toutes les catégo-ries d'information sont assujetties à un même processus descriptif et d'établissement des délais de conservation. Comme les archivistes passent en revue toute l'information liée à un système donné, ils n'ont pas à se demander si quelque chose leur a échappé. C'est donc dire que toute l'information stockée sur support imprimé ou informatisé, apparaît dans le contexte qui lui est propre.

*Cet article est le résumé d'un exposé de John McDonald, « An Approach to the Identification and Scheduling of EDP Data », présenté à la Society of American Archivists le 2 septembre 1984, à Washing-ton (D.C.).*





# Archives ordinolingués

BULLETIN

Vol. 3 - n° 2  
Été 1985

Canada

## Nouvelles acquisitions

Le fichier de données ordinolingués sur les résultats de l'élection fédérale de 1984 a été obtenu d'Élections Canada (anciennement Direction générale des élections). Ce fichier détaille les résultats par bureau de scrutin. Les fichiers de données des élections fédérales tenues en 1972, en 1974, en 1979 et en 1980 sont aussi disponibles.

## Calendriers de conservation des documents informatisés

Les Archives publiques du Canada (APC)

se fient depuis plusieurs années aux calendriers de conservation pour reconnaître, évaluer, acquérir et conserver les documents historiques créés par les ministères et organismes fédéraux.

Les calendriers de conservation sont des tableaux préparés par les gestionnaires de documents des organismes fédéraux pour indiquer combien de temps ils comptent garder leurs divers dossiers sur place à des fins administratives, puis combien de temps ils souhaitent les faire conserver par les centres de documents régionaux. Les organismes gouvernementaux fédéral pour que les Archives puissent déterminer la valeur historique des documents et, partant, choisir ceux qui seront transférés à la Direction des archives une fois les délais de conservation écoulés.

Les calendriers sont efficaces dans la mesure où chaque organisme a pris soin de bien structurer sa gestion des documents. L'importance de cette fonction vient d'être confirmée par l'adoption des lois sur l'accès à l'information et sur la protection des renseignements personnels et par la nouvelle politique du Conseil du Trésor sur la gestion des documents. Les calendriers de conservation ont assez bien fonctionné dans le cas des documents textuels ou imprimés, mais le processus laissait à désirer pour ce qui était de l'information enregistrée sur d'autres supports ou tenue par des secteurs organisationnels n'ayant aucun rapport avec la gestion des documents. Ainsi, les données ordinolingués ont souvent été laissées pour compte par les gestionnaires de documents de la plupart des organismes. Déjà dans les années 70 des efforts avaient été déployés pour repérer les organismes ordinolingués que possédaient les organismes fédéraux, mais il a fallu attendre la création

d'une Section des systèmes informatisés d'information à la Division des archives ordinolingués (DAO) pour qu'un programme systématique voie le jour. Étant donné les nombreux secteurs touchés par la question des calendriers de conservation, il a été décidé de privilégier le travail d'équipe. Outre les représentants des APC, on a jugé bon de regrouper un responsable des systèmes d'information, les gestionnaires de documents, le coordonnateur de l'accès à l'information et de la protection des renseignements personnels et tout autre intervenant direct. Des équipes de travail ont été mises à l'essai l'année dernière dans le cadre de trois projets-pilotes. Les modalités de description des données gérées à l'aide de systèmes automatisés (vue d'ensemble du système) ont été testées, et les résultats sont encourageants. On espère que tous les organismes fédéraux se conformeront à ces modalités au cours des prochaines années.

## Calendriers de conservation des documents informatisés : Vue d'ensemble du système

L'établissement d'un calendrier de conservation pour ce genre de documents est fondé sur une vue d'ensemble du système, qui sert à donner une idée générale du contenu de l'information et à fournir une représentation schématisée des grandes étapes du cheminement des données. Les organismes n'ont donc pas à décrire leurs disques ni leurs bandes en détail. Il leur suffit de faire connaître les applications qu'ils font de l'information et, à l'aide d'une approche pyramidale, de décrire les éléments-clés pouvant tirer profit des services des Archives publiques.

De nombreux systèmes sont conçus en fonction d'une seule tâche ou application (p. ex., sondage unique, impression d'étiquette, etc.). D'autres systèmes servent parfois à des applications multiples à partir du même ensemble de données (p. ex. une base de données utilisée à la fois dans le domaine du personnel et celui des finances). Par conséquent, la vue d'ensemble d'un système peut comprendre une ou plusieurs applications. Outre les détails administratifs tels que le nom de l'établissement et de l'unité, les éléments d'une vue d'ensemble sont les suivants :

**Titre du système :** Titre descriptif et significatif du système automatisé.

**Titre de l'application :** Description de chaque application permise par le système (un système peut appuyer une seule application ou, dans le cas de bases de données, plusieurs applications).

**But :** Description du but de l'application, en particulier la fonction ou les fonctions connexes qu'elle doit accomplir.

**Contenu de l'information :** Résumé du contenu des données traitées.

**Entrées/Sorties/Traitement/Sorties :** Aperçu général du cheminement des données au cours de leur traitement. Cette description ne se limite pas à l'entrée initiale des données ni à leur sortie finale. Elle est basée sur un « organigramme » qui explique comment les données sont traitées dans le système. La fréquence de certaines étapes est aussi mentionnée.

**Documentation/Renvois :** Description de tous les documents-clés (et de leur emplacement) : spécifications, procédures, codes et toute autre documentation d'appoint.

**Logiciel :** Désignation générale des programmes servant à gérer les données.

**Matériel :** Description du matériel utilisé pour élayer l'application.

**Événements marquants :** Résumé des grands événements qui ont influé sur la conception du système (année de production, années où ont eu lieu les grandes conversions du système et les modifications majeures aux structures d'enregistrement, etc.).

**Utilisateurs et conseillers techniques :** Noms, postes et titres des principaux utilisateurs et des personnes-ressources.

**Renvois au système de classification :** Mention du titre et du numéro du dossier dans le système de classification utilisé par l'organisme auquel les données traitées font allusion. Le système de classification des dossiers est tenu par le gestionnaire des documents de l'organisme.

**Renvois à la catégorie de dossiers :** Nom et numéro d'enregistrement de la catégorie de documents dont les données traitées font partie. Il s'agit d'une exigence de la Loi sur l'accès à l'information.

**Renvois au fichier de renseignements personnels :** Nom et numéro d'enregistrement du fichier de renseignements personnels auquel se rapportent les renseignements personnels traités par l'application. Il s'agit d'une exigence de la Loi sur la protection des renseignements personnels.